*J Neurophysiol* 97: 3517–3531, 2007.
First published March 14, 2007; doi:10.1152/jn.01066.2006.

# Development of Selectivity for Natural Sounds in the Songbird Auditory Forebrain

**Noopur Amin,[1] Allison Doupe,[3] and Frédéric E. Theunissen[1,2]**

[1]*Helen Wills Neuroscience Institute and* [2]*Psychology Department, University of California at Berkeley, Berkeley; and* [3]*Department of Physiology and Psychiatry, University of California at San Francisco, San Francisco, California*

Submitted ; accepted in final form 8 March 2007

**Amin N, Doupe A, Theunissen FE.** Development of selectivity for natural sounds in the songbird auditory forebrain. *J Neurophysiol* 97: 3517–3531, 2007. First published March 14, 2007; doi:10.1152/jn.01066.2006. In adult songbirds, auditory neurons in the primary auditory forebrain region of field L and a secondary auditory forebrain region of caudal mesopallium (CM) are highly responsive to natural sounds, such as conspecific song. Because these nuclei are involved in sensory representations of songs, we investigated how their function changes during development. We recorded neural responses to conspecific and tutor song and acoustically matched synthetic sounds in field L and lateral CM (CLM) of urethane-anesthetized juvenile male zebra finches of approximately 35 days of age. At this age, juvenile songbirds are memorizing the songs of their adult tutors but do not yet sing mature song. They are also starting to recognize songs of individual conspecifics. Compared with adult auditory forebrain neurons, juvenile neurons in field L were on average less responsive to auditory stimuli and exhibited less selectivity for natural sounds compared with the synthetic sounds. This developmental effect was more pronounced in the secondary subregions of L1 and L3 than in the primary thalamo-recipient subregion L2 of field L. CLM showed adultlike selectivity for natural sounds. Also, we did not find any evidence of memory for the tutor song in either field L or CLM. We note that the neural development of selective responses to conspecific song in the secondary subregions of field L is correlated with the emergence of individual song preference around 35 days of age. Therefore we suggest that the emergence of natural sound selectivity in field L could be important for the behavioral development of song recognition.

## INTRODUCTION

Perceptual discrimination of complex natural sounds is vital for the livelihood of animals in the wild: learning to distinguish the vocalizations of neighbors from strangers, of kin from nonkin, and of mates from nonmates is essential for identifying competitors and noncompetitors. Songbirds are a particularly attractive animal model to study the perception of complex vocalizations. Songbirds use song and other vocalizations for a myriad of communication tasks: males use song for territorial defense and mate attraction (Catchpole and Slater 1995); songs mediate pair-bonding and cooperation in male–female pairs (Hile et al. 2000; Marshall-Bal and Slater 2004); and communication calls produced by both male and female birds communicate basic needs for survival [i.e., contact calls are produced to maintain contact, separation calls to restore contact, begging calls to obtain food, and alarm calls to advertise danger (see Marler 2004)].

Because their survival depends on recognizing these conspecific sounds, birds are particularly proficient at such perceptual tasks. Birds can distinguish conspecific sounds even in dense acoustical environments in the wild (Appeltants et al. 2005; Aubin et al. 2000; Leonard and Horn 2005). In addition, earlier laboratory studies illustrated birds' ability to discriminate conspecific calls and songs from that of heterospecifics (Appeltants et al. 2005; Dooling et al. 1992). Songbirds were also shown to discriminate familiar from novel conspecific songs—they can discriminate the songs of neighbors from those of strangers, songs of relatives from those of nonrelatives, and mates from nonmates (Sherman et al. 1997). Thus early acoustical and social environments are crucial in shaping sound-based communication in the adult. Behavioral studies showed familiarity-based discrimination between different individual songs even in juvenile songbirds of both sexes (Clayton 1988; Riebel 2000; Riebel et al. 2002), an ability that is maintained and honed in adulthood (Miller 1979). Finally, the best-known and well-described example of the importance of early acoustical exposure in songbirds is that of a tutor song that is required for normal song production (Marler and Peters 1982).

The ascending auditory system of songbirds, where neurophysiological studies can be framed in the context of song learning behavior or in the more general context of recognition of complex communication sounds, has become a powerful model system to study the neural basis of the perceptual discrimination of complex and behaviorally relevant natural sounds (reviewed in Theunissen and Shaevitz 2006). Likely candidates for the neural basis of natural sound recognition include auditory forebrain regions such as field L, which is the first postthalamic processing stage, or secondary regions such as caudal mesopallium (CM) or caudomedial nidopallium (NCM) (reviewed in Bolhuis and Gahr 2006). These auditory forebrain regions reside between brain stem auditory processing centers and the high-level song system nuclei specialized for song production and learning (Nottebohm et al. 1976; Theunissen et al. 2004b) (and see Fig. 1).

In the discrimination of complex sounds, primary and secondary auditory forebrain areas were implicated in a series of experiments, all of which focused on adult songbirds. Electrophysiological studies (Grace et al. 2003; Hsu et al. 2004; Janata and Margoliash 1999; Leppelsack and Vogt 1976; Lewicki and Arthur 1996; Margoliash 1986; Muller and Leppelsack 1985; Stripling et al. 1997; Woolley et al. 2005), immediate early

Address for reprint requests and other correspondence: F. Theunissen, University of California, Berkeley, Psychology Department, 3210 Tolman Hall, MC #1650, Berkeley, CA 94720–1650 (E-mail: theunissen@berkeley.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

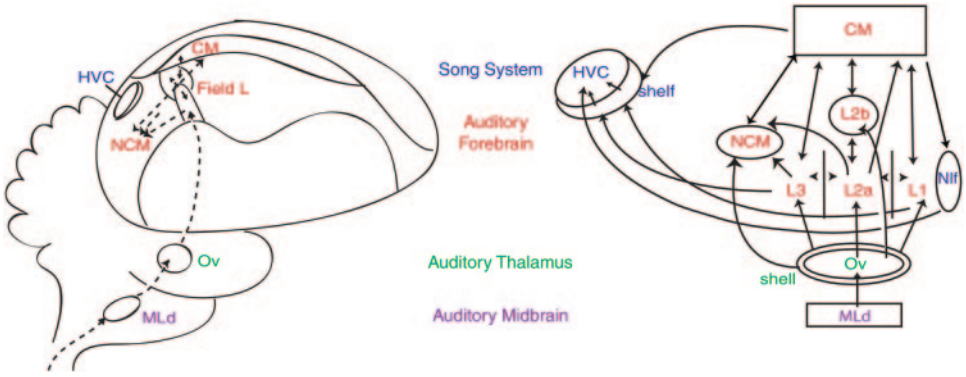



FIG. 1. Schematized depiction of a sagittal section of the male zebra finch brain (*left*) and anatomical subdivisions and connectivity of the avian auditory forebrain in relation to the song system (*right*). MLd, mesencephalicus lateralis dorsalis; Ov, ovoidalis; NIf, nucleus interface of the nidopallium; NCM, caudal medial nidopallium; CM, caudal mesopallium; HVC, high vocal center.

gene studies (Jarvis and Nottebohm 1997; Mello and Clayton 1994; Mello and Ribeiro 1998; Mello et al. 1992; Ribeiro et al. 1998), and lesion studies (MacDougall-Shackleton et al. 1998) showed that natural sounds, such as conspecific song or calls, are robustly and preferentially represented in these areas. NCM and CM were also previously implicated in the neural recognition of familiar songs (Chew et al. 1995; Gentner and Margoliash 2003; Mello and Clayton 1994), including the tutor song (Phan et al. 2006; Terpstra et al. 2004).

We showed in a previous study that field L and lateral caudal mesopallium (CLM) of adult male zebra finches respond selectively to natural sounds compared with a specific set of synthetic sounds that are matched statistically to natural sounds (Grace et al. 2003). We suggested that this neural selectivity might be useful for discriminating behaviorally relevant natural sounds. In this study, we examined the neural selectivity in these same brain areas at a critical point during development: we studied 35-day-old male juveniles because young male zebra finches can learn to produce a good copy of a tutor song to which they have been exposed until they are 35 days old (Böhner 1990). Moreover, zebra finches of both sexes are beginning to show preference for individual conspecific songs around this time (Clayton 1988). These behavioral observations suggest that the auditory discrimination of young birds at 35 days must be approaching maturity or potentially be similar to that of adults.

On the other hand, these young birds have not yet approached sexual maturity and maturational changes are known to occur in other forebrain regions throughout development (Heinrich et al. 2002; Iyengar et al. 1999; Nordeen and Nordeen 2004). In addition, young males have not yet learned to produce a mature song. If vocal experience (i.e., learning to sing) plays a key role in shaping selectivity in auditory forebrain regions, we would expect to find little selectivity. This is the case in the song system, which receives afferent input from the auditory forebrain: song system auditory responses recorded around this time are substantially less selective than those found in the adult (Doupe 1997; Solis and Doupe 1997) and, more specifically, follow the vocalization that the bird is currently producing (Nick and Konishi 2005b). It is therefore difficult to make a strong hypothesis about what might be expected in the auditory telencephalic regions of field L and CLM; the outcome of our study will set the stage for understanding the relative contribution of perceptual and vocal experience, as well as maturation, on the development of the auditory selectivity for natural sounds observed in the adult. In addition, in the context of song learning behavior, we were interested in whether we would find selective responses for the tutor song in the juvenile auditory forebrain.

## METHODS

### *Experimental procedures to obtain audiograms*

To measure the overall maturation of the lower auditory system, we obtained audiograms from auditory brain stem responses. All animal procedures were approved by the Animal Care and Use Committee at University of California Berkeley and University of California San Francisco (UCSF). Audiograms are functions relating hearing thresholds to frequency. We obtained audiograms from urethane anesthetized zebra finches at posthatch day 10 (PHD 10), PHD 20, and in adults (PHD $>$ 90). These birds were not used in any of the electrophysiological recordings. For the experiment, the bird was placed on a bird sling inside a double-chambered sound-proof box. The bird's head was carefully taped onto the sling to prevent any movement. Free-field sounds were played from a speaker placed directly in front of the bird at 15 cm. Before each experiment, the sound delivery system was calibrated using a *Ban* dK microphone. Three adult males and one female ($n = 4$) were used for the adult data; three 20-day-old females ($n = 3$), two 10-day-old males, and two 10-day-old females ($n = 4$) were used for the juvenile data. The threshold levels for males and female birds were statistically indistinguishable in the adult [two-way ANOVA with gender and frequency as factors: $F(1,43) = 0.39$, $P > 0.5$] or when the adult and 20-day-old data were combined [three-way ANOVA with gender, frequency, and age as factors: $F(1,76) = 0.39$, $P > 0.5$]. The data from males and females were therefore grouped together. It should be noted, however, that our sample is not large enough to detect any potential small differences.

Audiograms were obtained by determining the sound intensity threshold level of auditory-evoked potentials in the brain stem [also known as the auditory brain stem response (ABR)]. The evoked potentials were recorded using low-impedance pin electrodes. A recording electrode was placed into the cerebellum just above the auditory brain stem. A second, differential electrode was placed in the forebrain. One hundred to 200 response waveforms to 20-ms pure tones (rise/fall time of 5 ms) of frequencies between 400 and 10,000 Hz were obtained for a range of sound levels measured in 5-dB increments. The threshold was found by comparing peak values in the averaged ABR waveform during sound to the distribution peak amplitudes in the averaged waveform obtained in the 40 ms preceding the stimulation. The threshold value was defined to be the lowest sound level that gave rise to peak max or min values in the ABR response with a $P$ value of $<0.001$, in comparison to background.

The audiogram recordings were performed in the laboratory of A. Doupe at UCSF, and these procedures were approved by the ACUC committee of that institution.

### *Animal procedures for acute neurophysiological recordings*

Young male zebra finches (*Taenopygia guttata*) of roughly PHD 35 days were used in all of the subsequent neurophysiological experiments (mean = 36.9 days, SD = 3.7 days). The birds were anesthetized either 2 days before the acute recording or on the day of the physiological recordings. In the prior case, about 0.03 ml Equithesin (0.85 g of chloral hydrate, 0.21 g of pentobarbital, 0.42 g of $MgSO_4$, 8.6 ml of propylene glycol, 2.2 ml of 100% ethanol, to a total volume of 20 ml with $H_2O$) was administered intramuscularly (im). In the latter case, the bird was sedated by a combination of Metofane (Mallinckrodt Veterinary, Mundelein, IL) delivered to the respiratory system and about 25–30 $\mu$l of 20% urethane delivered im. Equithesin is known to depress activity in the forebrain for substantial periods of time and thus the birds given this anesthetic were allowed to recover for 2 days. The surgery itself consisted of the following: removing a small region of skin on top of the skull after having immobilized the bird with the aid of a beak holder and ear bars on a stereotaxic apparatus; removing a 1-mm square of the top layer of skull around the field L reference; marking the reference (1.5 mm lateral and 1.2 mm rostral from the midsagittal y-sinus) with ink to guide electrode penetrations; and, finally, gluing a stainless steel post onto the skull with dental cement.

For those birds that were surgically fitted with the steel post 2 days before the recording, the bird was then anesthetized with three im injections of 25–30 $\mu$l each of 20% urethane administered at 0.5-h intervals on the recording day. Those birds that were administered one dose of urethane for the surgery then received an additional two injections of 25–30 $\mu$l each of 20% urethane at 0.5-h intervals after surgery. For the latter case, there was a 1.5- to 2-h lag between the end of the surgical procedure and the start of the recording session, thereby effectively eliminating any residual effects of Metofane on the bird's nervous system. It should be noted that overall the birds received variable amounts of urethane, from 75 to 90 $\mu$l (mean urethane use for all birds = 83.3 $\mu$l; SD = 6.2 $\mu$l), depending on the individual bird's level of wakefulness and activity. However, these differences did not have a significant effect on the level of neural responses. A correlation analysis between the amount of urethane administered and auditory activity measured as a z-score in response to tutor song (see following text) had $R^2 = 0.007$, and a regression analysis showed that the slope of $-0.013$ was not significantly different from a slope of zero ($P = 0.36$).

It should be noted that urethane anesthesia depresses the spontaneous activity of auditory forebrain neurons (Capsius and Leppelsack 1996), although the overall effects of this anesthetic on the auditory forebrain (Capsius and Leppelsack 1996; George et al. 2004) are not as pronounced as those observed in the song system (Cardin and Schmidt 2004). Although it will be useful to obtain data in future experiments from awake restrained or chronically implanted juvenile birds, in the current study we were focused on a comparison of juvenile birds to our prior studies of urethane-anesthetized adults (Grace et al. 2003) and thus used the same conditions, including anesthesia.

To set up for the recording session, the bird's head was immobilized by attaching the steel post to the stereotax frame. The lower layer of the skull and the dura were removed around the ink-marked field L reference. Extracellular tungsten electrodes (1- to 4-M$\Omega$ resistance) were lowered into the brain using a microdrive. The bird and the stereotax were then placed in a calibrated, double-walled anechoic chamber where a speaker was used to present the stimuli. The volume of the speaker was set to deliver zebra finch songs at either of the following intensities: peak levels at 70 or 85 dB SPL (*Ban* dK Sound Level Meter, RMS weighting type B). The speaker was placed 20 cm in front of the bird's head. All animal procedures were approved by Animal Care and Use Committee at UC Berkeley.

### *Stimulus design and stimulus recordings*

The stimuli included natural and synthetic sounds. The natural sounds were *1*) songs of unrelated adult male zebra finches or conspecific song (Con); *2*) the conspecific song played in reverse (CRev); *3*) the reverse-order version of Con (CRO), in which the sequence of syllables is played in reverse while the temporal order within each syllable is maintained; *4*) the bird's tutor's song (Tutor), which in these experiments was always the bird's father; *5*) the tutor song played in reverse (TRev); and *6*) the reverse order version of the tutor song (TRO). The synthetic sounds were: *1*) a sequence of random pure tones (Pips); *2*) compound pure tones (Tones) in which 20 samples from the random pips group were added together; and *3*) broadband white noise (WN). The synthetic sound ensembles were designed to match several statistical parameters found in the zebra finch song, as described later. The *top panels* of Fig. 3 show spectrograms of particular exemplars of a subset of the stimulus classes used in our experiments.

We used 20 Pips, 20 Tones, and 20 WN stimuli in all, each of which lasted exactly 2 s. To compare neural responses to the synthetic stimuli to those to song, we matched certain acoustic parameters. The power spectrum distribution of zebra finch song was randomly sampled to obtain frequencies of pure tones and the Pips sound consisted of a succession of such pure tones in time and with the same overall power spectrum as that of conspecific song. The length of the tone pips and the interpip silences were drawn from a Gaussian distribution that approximated the distribution of the length of song syllables [95 $\pm$ 66 (SD) ms] and intersyllable silences (37 $\pm$ 25 ms). The onset and offset ramp of each tone pip was a 25-ms cosine function, loosely matching the amplitude envelope of song syllables. The intensity of the individual tones was randomly varied uniformly over a logarithmic scale, although the overall power of the Pips sound was matched to that of the conspecific songs. The Pips stimulus could therefore be considered to be the best possible descriptor of zebra finch song made with simple tone pips.

The Tones ensemble is the broadband extension of the Pips ensemble. The Tones were synthesized by adding 20 different Pips sounds together and normalizing the result to retain the same overall power as song. Thus the spectral composition of the Tones stimuli was random within the power spectral distribution of zebra finch song and, although the overall amplitude envelope no longer matched the statistics of song, the distribution of intensities in any narrow-frequency band was similar to that found in song.

Finally, the WN samples were band limited between 250 Hz and 8 kHz, spanning the audible range of zebra finches. The power spectrum was flat between these boundaries and the level was set so that the overall power summed across frequencies was matched with that of song, Pips, and Tones.

Spectrographic representations of these matched synthetic stimuli can be seen in the *top panels* of Fig. 3. A detailed description of these stimuli and their acoustical properties can be found in Grace et al. (2003). In that paper, we compared the responses of auditory forebrain neurons (in field L and CLM) to song and to synthetic stimuli to obtain a measure of neural selectivity for the higher-order statistical properties of conspecific song. The same analysis was performed here on data from juvenile birds, thereby allowing for a direct comparison of the adult and juvenile studies.

The song ensemble consisted of the tutor and conspecific songs and temporally manipulated versions of those songs. To obtain recordings of these songs, each bird was placed in a sound-attenuated box for about 5–10 h and exemplars of the bird's song were recorded. All song bouts were listened to and viewed as spectrograms and the most frequently sung version was chosen as the song stimulus. Although the duration of the songs varied, the mean duration of songs was similar to that of the synthetic stimuli: the duration of the set of conspecific songs was 2.08 s (SD 0.63 s) and that of the tutor songs 2.18 s (SD 0.36 s). The songs were then digitized using TDT2

hardware and reversed using custom software. The Con stimulus set consisted of songs from 20 unrelated adult male zebra finches. It was recently shown that spectral and temporal patterns occurring in zebra finch song are well represented in 20 songs such that the sampling error in both the power spectrum and the modulation spectrum is minimal (Singh and Theunissen 2003). The juvenile's father's song was taken to be its Tutor song. All young males in our colony were raised in cages only with their parents and clutch-mates, thereby increasing the chances that the juvenile will learn the song of his father: zebra finches learn the song to which they are most socially exposed (Slater et al. 1988). In addition, visual barriers were placed between all cages to assist with song learning from the tutor bird and to prevent the young males from learning the song of other conspecific males (Eales 1985). Because the birds used in our acute neurophysiological recordings were killed at the end of the experiment, we were unable to track their song learning and we did not determine whether individual experimental birds had entered the subsong stage of song learning. However, to monitor song learning in our overall colony, we tracked the song development of other young birds from 40 days of age to adulthood (6 families, 39 birds) and compared their song to that of their presumed tutor. Although there was some variability across families, song learning from the father was robust (see Amin et al. 2004).

All songs were filtered between 250 and 8,000 Hz and then normalized to have identical overall mean power during the nonsilent parts of the song. We compared neural responses at two sound levels that are within the behavioral range. Each experimental bird was played songs at only one of two sound levels: at either peak levels of 70 dB SPL (for 7 birds, 44 stimulus-excited sites) or 85 dB SPL (for five birds, 25 stimulus-excited sites). Average z-scores, a measure for response strength (see following text for details), were calculated for all stimulus-excited sites in response to our song and synthetic stimulus ensembles. A comparison between responses to sounds played at the two intensities revealed that there was no difference in the average responses to these two sound levels [$t(737) = 1.39$, $P = 0.16$]. Similar observations were made for sensorimotor neurons in the song system in response to song (Margoliash 1986). In subsequent analyses, responses to the two sound levels were thus grouped.

At each recording site, the search stimuli (White Noise and either Con or Tutor) were played to determine whether the site was responsive. If the response to *either* of the two search stimuli was significantly different from the spontaneous firing rate, as determined by an on-line *t*-test, then the entire stimulus ensemble was presented to that recording site. The song stimuli consisted of: 10 (and sometimes 20) presentations of Tutor, TRev, and TRO; 10 presentations each of three different Cons, and one of those Cons' Reverse (CRev) and Reverse Order (CRO) versions. The synthetic stimuli included 10 presentations each of three different Pips, three different Tones, and two different WN sounds. The stimulus presentation order was randomized per trial and a random interstimulus interval with a uniform distribution of 7 to 8 s was used. Finally, two seconds of spontaneous spiking were recorded both before and after the stimulus presentation.

### *Electrophysiology and experimental protocol*

One to three electrode penetrations were achieved per bird and the auditory forebrain regions of field L and CLM were systematically sampled every 75–125 microns per penetration to estimate the number of responsive versus nonresponsive sites. Penetrations spanned the mediolateral (from 1.2 to 1.8 mm lateral from the y-sinus) and the rostrocaudal (from 1.2 to 1.8 mm rostral to the y-sinus) axes of the auditory forebrain. To be consistent with the adult recording sites, we aimed for similar electrode penetrations in the juveniles, systematically sampling along both the mediolateral and rostrocaudal axes. For each recording site, the spike arrival times were recorded by thresholding the extracellular voltage trace. Both single and multiunits (defined as a small cluster of two to five neurons based on spike shapes) were recorded to be able to directly compare the juvenile data with those of adults recorded with a similar experimental protocol (see Amin et al. 2004; Grace et al. 2003). The data from single units and multiunits were analyzed separately. Because similar conclusions were obtained in both cases (see RESULTS), we were able to combine the data for some of the analyses, to increase statistical power.

For birds in which more than one electrode pass was achieved, only one electrolytic lesion (100 $\mu$A for 5 s) was made at the end of all but the final pass to allow for future recording site reconstruction and was made 300–400 microns subsequent to the last recording site's depth. In the final electrode penetration, two lesions 300 microns apart were created for calibrating our depth measures (see following text). Lesions were generally created outside of the auditory forebrain and no differences in response properties were observed between recordings before and after the lesions.

### *Histology and anatomical reconstructions*

After recording, birds were overdosed with Equithesin and transcardially perfused with 0.9% saline, followed by 3.7% formaldehyde (10% formalin) in 0.025 M phosphate buffer. The brain was sunk in 30% sucrose and 3.7% formaldehyde to prepare it for histological procedures. The brain was sliced parasagittally in 40-$\mu$m-thick sections using a freezing microtome. Alternating brain sections were stained with both cresyl violet and silver stain, which were then used to visualize electrode tracks and electrolytic lesions.

Recording site reconstruction involved measuring both the distance from the entry of the electrode pass to the lesion and the distance between successive lesions and comparing these distances in microns with the reading of our independently calibrated microdrive used during the experiment. The sites were then reconstructed with the aid of the experimental log, containing microdrive-measured distances between subsequent sites, as a reference. Using well-known anatomical landmarks such as the pallial-subpallial lamina (LPS) and differences in cell size, shape, and density as described in the literature (Fortune and Margoliash 1992), neural sites were then assigned to either CLM, or thalamo-recipient subdivision L2 (L2a or L2b), or subregions L1 and L3. L2a and L2b were the most readily distinguishable subregions based on cell shape and size. Region L1 was defined as the area that was dorsal to the boundary of L2 and below the lamina that divides the nidopallium and the mesopallium. Region L3 was defined as the area below the ventral boundary of L2 within the nidopallium. We were not able to distinguish a boundary between subfield L3 and subfield L and all the ventral recording sites were assigned to L3 with that caveat in mind. Because units in L1 and L3 had statistically similar properties in this study, they were grouped together and labeled L1/L3. Based on the stereotaxic measurements of our electrode penetrations (1.2–1.8 mm from the midline), we believe that all of our CM recordings were from lateral CM (Vates et al. 1996).

### *Data analysis*

A unit was defined as responsive if its average firing rate to *either* WN or Con was significantly different from its spontaneous firing rate ($P < 0.05$, two-tailed paired *t*-test). WN and Con were used to determine responsive units for the following reasons: *1*) WN is a nonspecific broadband stimulus that is often used to quickly characterize auditory units; *2*) conspecific song elicits robust responses in the auditory forebrain regions of field L and CM of adult zebra finches (Amin et al. 2004), and thus we wanted to use this preferred stimulus to probe the juvenile songbird auditory forebrain; and *3*) we knew from the adult studies that it is rare for field L neurons not to respond to Con and WN, and yet respond to other sounds. Although Tutor song was sometimes used as a search stimulus because of its behavioral relevance for the songbird at this age, we did not use it in defining an auditory unit in our post hoc analysis because we found no

differences in response to Tutor or Con. Thus we kept with the more general version of song when defining a unit as responsive.

For each stimulus, all responsive sites were given a z-score, which characterizes the normalized difference between the stimulus-evoked mean firing rate and that of the two second background activity preceding the stimulus. The z-score is calculated as follows

$$z = \frac{\mu_S - \mu_{BG}}{\sqrt{\sigma_S^2 + \sigma_{BG}^2 - 2\text{Covar}(S, BG)}} \quad (1)$$

where $\mu_S$ is the mean response during the stimulus, $\mu_{BG}$ is the mean response during the background, $\sigma_S^2$ is the variance of the response during the stimulus, and $\sigma_{BG}^2$ the variance of the response during baseline. In calculating a neuron's z-score to a particular stimulus type, responses of that neuron were averaged to all presentations for that particular stimulus type. For instance, a site's response to three exemplars of Con was averaged together when calculating that site's single z-score measure to Conspecific song.

All significant responsive sites were classified as stimulus excited or stimulus inhibited. A stimulus-excited site was defined to have a positive z-score value for Con and, likewise, a stimulus-inhibited site had to have a negative z-score in its response to Con.

To quantify the difference in neural responses to the stimulus classes, we used the psychophysical d′ measure, which is used to quantify neural selectivity in the avian song system and auditory forebrain (Theunissen et al. 2004a,b). The d′ measure for the response difference between two stimuli A and B is calculated as follows

$$d'_{A-B} = \frac{2(\mu_A - \mu_B)}{\sqrt{\sigma_A^2 + \sigma_B^2}} \quad (2)$$

where $\mu_A$ and $\mu_B$ are the mean responses to stimulus A and B, respectively, and $\sigma^2$ is the variance of the response. For positive d′ values, stimulus A evoked a greater response, and for negative d′ values, stimulus B evoked a greater response. Values of d′ close to 0 indicate no difference in the mean responses evoked by the two stimuli. The d′ measure will also give negative values when stimulus A elicited a greater suppression than the suppression obtained by stimulus B. Thus it is necessary to distinguish the stimulus-excited from the stimulus-inhibited sites when interpreting d′ values for selectivity purposes. For any given unit, a d′ value was calculated for every pairwise comparison, taking into account individual exemplars of a stimulus class, and then those values were averaged into a single d′ value for that unit's selectivity to the two stimuli. For example, for most units we obtained responses to one exemplar of the Tutor stimulus and three exemplars of Conspecific songs, yielding three d′ measures for the Tutor–Con comparison for that unit, which are then averaged to one d′ value per unit. It should be noted that our descriptions of neural responses and the corresponding quantification with z-scores and d′ measures depend solely on the mean firing rates and ignore potential stimulus-dependent information that could be present in systematically different spike patterns.

Two additional post hoc analyses addressed whether responsivity and selectivity are correlated during development. First, we calculated the correlation coefficient between responsivity, given by the z-score, and selectivity, given by the d′ value. Second, we compared the juvenile selectivity to the adult selectivity of subsets of adult neurons that were chosen to have the same distribution of responsivity as that of the young neurons. To investigate this issue, the adult z-score distribution for Con (adapted from Grace et al. 2003) was resampled to match the juveniles' z-score distribution for Con. This resampled adult distribution was created by randomly choosing and rejecting adult neurons such that the resampled set had the same count histogram of z-scores as that of the juvenile distribution: i.e., adult cells were randomly chosen until each count bin for a given z-score interval had the same number of cells as that of the juvenile (see Fig. 7*B*). Ten such resampled adult z-score distributions were created. The mean d′ value for the Con–Pips comparison was obtained for each of these 10 resampled distributions, from which an average d′ value was calculated. If the new resampled adult average d′ value is no different from that of the juveniles', then we would conclude that responsivity and selectivity coemerge during auditory development. However, if the resampled adult d′ distribution is still greater than that of the juveniles', then that would imply that the difference in selectivity cannot solely be explained by differences in the strength of each response. We limited this analysis to subregions L1 and L3 where responsivity and selectivity were still immature in the juveniles compared with the adults (more details provide later).

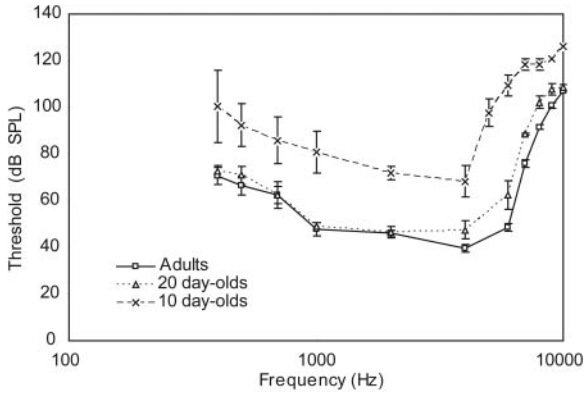



FIG. 2. Audiograms of posthatch day (PHD) 10, PHD 20, and adult zebra finches, using the auditory evoked potentials in the brain stem. These audiograms show that by PHD 20, the auditory brain stem is already functioning as in adults.

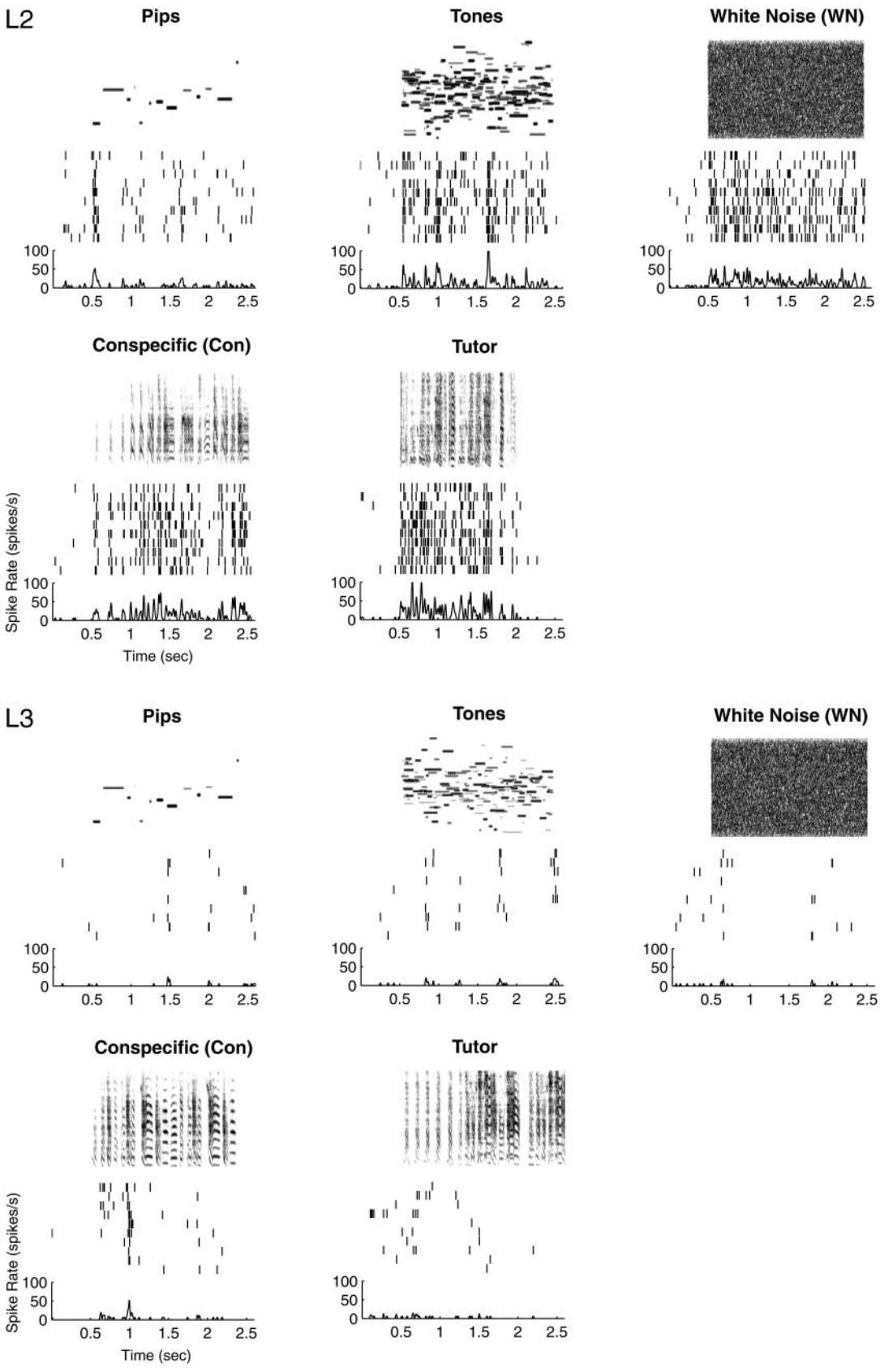
A final analysis looked at the coefficient of variation (CV) of juvenile mean firing rates across trials compared with that of the adults. The CV is the ratio of the SD of the background-subtracted evoked spike rate and the mean of this background-subtracted rate. The SD and mean are estimated over the 10 spike trials that are obtained for each stimulus. The CV measures the neural variability of the response in normalized units. We computed the CV for both juvenile and adult stimulus-excited cells. To obtain reliable measures, the CV was calculated only for neurons that had a background-subtracted spike rate >1 spike/s in response to Con.

## RESULTS

The principal goal of the study was to investigate how natural sound selectivity in the avian auditory telencephalic regions of field L and CLM changes during juvenile development. We previously showed that field L and CLM neurons respond preferentially to conspecific song over matched synthetic sounds (Grace et al. 2003). Here we extend the previous work to address whether similar natural sound selectivity exists in juvenile zebra finches. A second goal was to assess the role of the tutor song in the neural responses of these developing forebrain regions.

### *Audiogram results*

Before studying the selectivity of auditory forebrain responses in young animals, we wanted to ensure that zebra finches, known to be altricial animals, have normal hearing at this age. For this purpose, audiograms of PHD 10, PHD 20, and adult zebra finches were obtained by measuring auditory-evoked potentials in the brain stem (see Fig. 2). The audiograms that we measured in the adult bird were similar to those found by other researchers in zebra finches (Zevin et al. 2004), in bengalese finches (Woolley and Rubel 1999), and in bud-



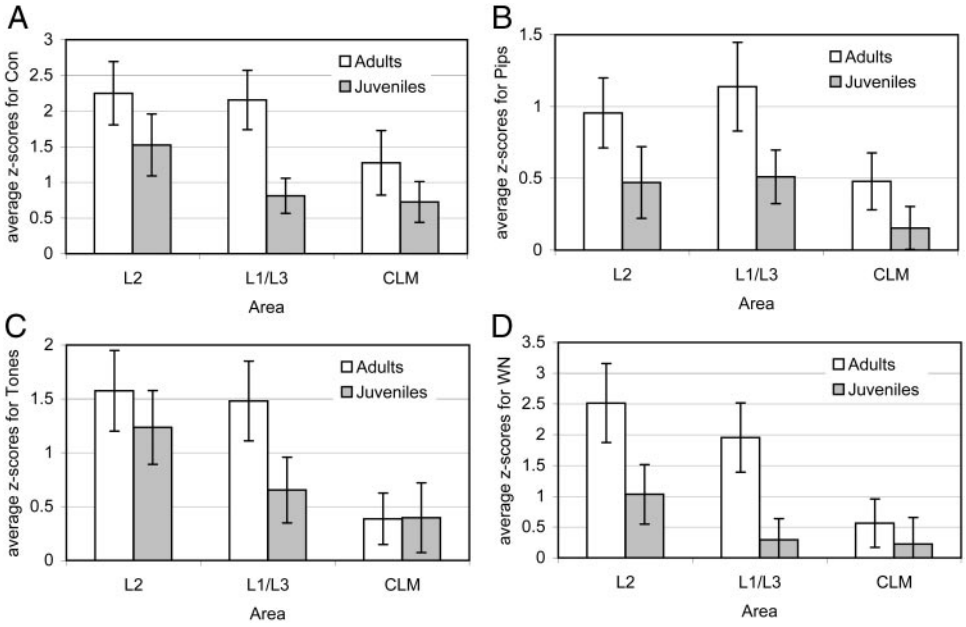



FIG. 4. Comparison of juvenile and adult mean z-scores across areas [L2, L1/L3, and lateral caudal mesopallium (CLM)] for the following stimuli: *A*: Con. *B*: Pips. *C*: compound pure tones (Tones). *D*: broadband white noise (WN). L1/L3 subareas show the greatest developmental effect. Error bars represent 2 SEs.

gerigars (Brittan-Powell and Dooling 2002). The audiograms of PHD 10 birds show significantly higher thresholds than those in adult birds, implying that the lower levels of the auditory system are not fully mature by that date. The audiograms of PHD 20 birds, on the other hand, are very similar to the audiograms obtained from adults, indicating that the sensitivity of the auditory brain stem of 20-day-old zebra finches is mostly adultlike, with only small differences in the higher-frequency range.

### *Auditory responses*

Having determined that by 20 days zebra finches have adultlike auditory thresholds, we examined the selectivity of auditory responses in the auditory forebrain regions of field L and CLM in birds at approximately 35 days of age (mean = 36.9 days, SD = 3.7 days). We obtained detailed neurophysiological data from 114 auditory forebrain units (from both field L and CLM) from 12 urethane-anesthetized juveniles. Of these units, 88 were classified as responsive because responses to Con or WN were significantly different from spontaneous activity. The percentage of responsive units was 77% (88/114) but it should be noted that this number does not reflect the actual percentage of auditory neurons in the sampled brain regions because sampling was biased by our search protocol where we actively searched for neurons that appear to respond to sounds. For that same reason, the number cannot be compared with the number that we reported in the adults (54%), which was obtained by systematically sampling at 100-micron intervals. Moreover, as we will show later, the neurons that were auditory had smaller response strengths.

Of the 88 responsive sites from field L and CLM, 69 were classified as stimulus-excited sites where the z-score was $>0$ and 19 were classified as stimulus-inhibited where the z-score was negative. Because the sample size of inhibited sites was small, we focused on the stimulus-excited sites for the selectivity analyses. We assigned 62 of the excitatory units to one of three main anatomical regions: 25 of the stimulus-excited units were in L2 (L2a and L2b together); 24 sites were in other subregions of field L (L1 and L3); and 13 sites were in CLM. Seven units were not classified because the histology was ambiguous. Figure 3 shows examples of two recording sites to a subset of the stimuli. The *top panels* show example responses from one of the more responsive cells, which were more typically found in subregion L2 (as was also the case for this example). The *bottom panels* show responses from one of the less-responsive sites from the data set, which were more typically found in subregion L1 or L3 (this particular neuron was from region L3). These example cells are labeled in the scatterplot of z-scores and d′ values in Fig. 8*A* to illustrate where they lie in the distribution of all responsive cells.

### *Responsivity in adult versus juvenile birds*

We examined the stimulus-excited sites (single- and multineuronal units combined) in both field L and CLM to test whether juvenile responsivity and selectivity would be similar to those in adults. Figure 4, *A–D* shows the average z-scores responses for L2, L1/L3, and CLM in response to Con, Pips, Tones, and WN, respectively, for both juveniles and adults (adapted from Grace et al. 2003). We performed a three-way ANOVA to investigate the effects of age, stimulus type, and brain region on z-score responses. The analysis shows a main effect for age [$F(1,907) = 39.52$, $P < 0.0001$]: the responses in juveniles are clearly depressed relative to the responses in the adults. The analysis also showed a main effect for stimulus type [$F(3,907) = 8.65$, $P < 0.0001$]: both adult and juvenile

FIG. 3. Spectrographic representation (frequencies ranging from 500 to 8,000 Hz on the *y*-axis and time in seconds on the *x*-axis) of exemplars of a subset of stimulus types used in our study and corresponding neural responses for 2 different sites. Note that sound begins at 0.5 s. For the neural response, both the spike raster for 10 trials (*middle*) and the mean response (denoted by spikes/s on the *bottom*) are shown. Recording site (*top*) is from the L2a region [z-score to conspecific song (Con): 1.40; d′ for conspecific song–random pure tones (Con–Pips): 3.68], whereas the recording site in the *bottom panels* is from L3 (z-score to Con: 0.83; d′ for Con–Pips: 1.17). Each of these neurons is represented with an arrowhead on the scatterplots of z-score and d′ values in Fig. 8*A*.

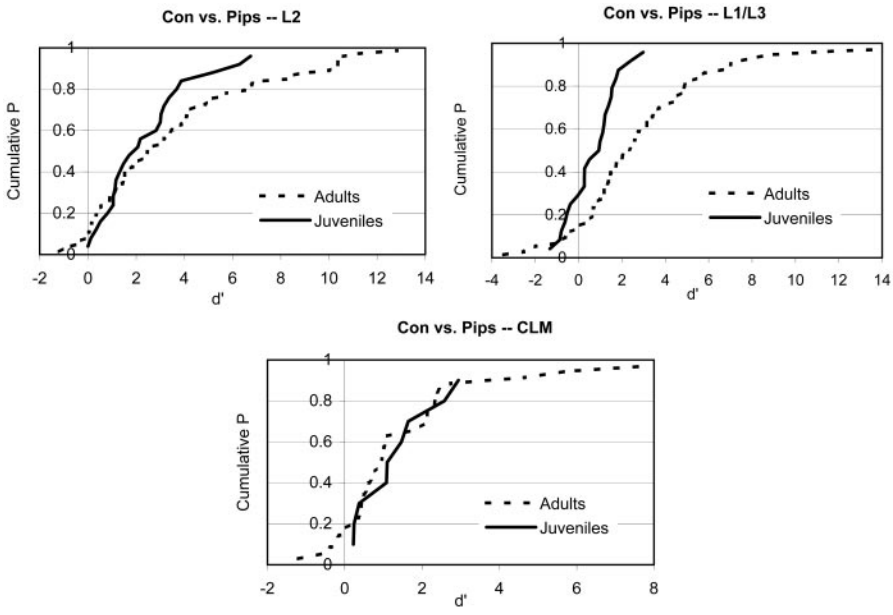



FIG. 5. Cumulative distributions of the d′ values for the Con–Pips comparison across the subareas for both juveniles and adults. L1/L3 subregions show the greatest divergence in the cumulative curves and thus the greatest developmental effect.

responses are greatest to Con and least to Pips. Finally, the analysis also showed a main effect for brain region [$F(2,907) = 19.28$, $P < 0.0001$]: responses are smaller in L3/L1 and CM than in L2. We also tested for the statistical significance of the three possible two-way interactions. We found a significant interaction between age and brain region [$F(2,907) = 3.52$, $P = 0.03$]: the difference in responsivity between adults and juveniles is the greatest in L1/L3 (the response in juveniles is only 33.7% of that in adults across all the four stimuli) and much smaller in L2 (58.8%) and CLM (51.7%). We also found a significant interaction between age and stimulus type [$F(3,907) = 3.31$, $P = 0.0195$]: for Con, the response in juveniles is 44.0% of that in adults, 44.0% for Pips, 63.1% for Tones, and 28.6% for WN. These differences will be further analyzed in the more appropriate pairwise comparisons provided by d′ values in the selectivity analysis (see subsequent section).

### *Selectivity for natural versus synthetic sounds*

To quantify the selectivity for conspecific song, we calculated a d′ for each pairwise comparison and for each site. Figure 5 shows cumulative probability distributions for d′ values for combined single and multiunits for both the juvenile and adult Con–Pip comparisons in all three brain regions. The juvenile Con–Pips distribution is shifted to the right relative to zero, indicating selectivity for Con relative to Pips. However, when compared with the adult distributions, the juvenile distributions in field L (both L2 and L1/L3) are shifted to the left, illustrating the fact that the juvenile auditory forebrain region of field L does not exhibit the selectivity found in the adult. In particular, the units with the highest degree of selectivity appear to be absent in the young birds. In addition, the differences in selectivity between the juveniles and adults appear to be greater in L1/L3 than in L2.

A quantitative analysis of the differences in the means of these distributions for both the Con–Pip and the other Con–Synthetic comparisons is shown in Fig. 6. The bar graphs in Fig. 6 show that the average differences in selectivity for conspecific song between adult and juveniles are absent in CLM, small in L2, and large in subareas L1 and L3. Even though there is some selectivity for Con over Pips in juvenile L1 and L3, this selectivity is very limited compared with that of the adults (Con–Pips juvenile mean d′ for L2 = 2.33, $t = 6.19$, and $P < 0.001$; Con–Pips juvenile mean d′ for L1/L3 = 0.67, $t = 2.88$, and $P = 0.008$; Con–Pips juvenile mean d′ for CLM = 1.26, $t = 4.30$, and $P = 0.002$). Similar differences in the degree of selectivity is found for other comparisons, where sites in juvenile L2 and CLM responded more to Con than to the other synthetic stimuli, as compared with L1/L3 (Con–Tones juvenile mean d′ for L2 = 0.52, $t = 1.87$, and $P = 0.07$; Con–Tones juvenile mean d′ for L1/L3 = 0.03, $t = 0.10$, and $P = 0.92$; Con–Tones juvenile mean d′ for CLM = 0.54, $t = 2.21$ and $P = 0.05$; Con–WN juvenile mean d′ for L2 = 0.87, $t = 2.83$, and $P = 0.009$; Con–WN juvenile mean d′ for rest of field L = 0.33, $t = 0.87$, and $P = 0.39$; Con–WN juvenile mean d′ for CLM = 0.43, $t = 1.26$, and $P = 0.23$).

It should also be noted that the only mean d′ value for juveniles that is highly significantly different from zero is the one for the Con–Pips comparison [Con–Pips juvenile mean d′ (for combined single- and multiunit data in all of field L) = 1.51, $t = 6.09$, and $P < 0.001$; Con–Tones juvenile mean d′ = 0.28, $t = 1.39$, and $P = 0.16$; Con–WN juvenile mean d′ = 0.49, $t = 2.19$, and $P = 0.03$]. Single- and multiunit analyses also showed similar trends: Con–Pips juvenile mean d′ (for single units) = 1.71, $t = 4.62$, and $P < 0.001$ and Con–Pips juvenile mean d′ (for multiunits) = 1.31, $t = 3.93$, and $P < 0.001$; Con–Tones juvenile mean d′ (for single units) = 0.34, $t = 1.14$, and $P = 0.26$ and Con–Tones juvenile mean d′ (for multiunits) = 0.22, $t = 0.79$, and $P = 0.43$; Con–WN juvenile mean d′ (for single units) = 0.48, $t = 1.53$, and $P = 0.13$ and Con–WN juvenile mean d′ (for multiunits) = 0.50, $t = 1.53$, and $P = 0.13$. Even though the mean responses to Con and WN are very similar in juvenile and adults birds, the temporal response profile to the two sounds can be quite different (as shown in Grace et al. 2003). Indeed, information theoretic analysis shows that conspecific song elicits higher information rates than do broadband noise stimuli in adult birds (Hsu et al.

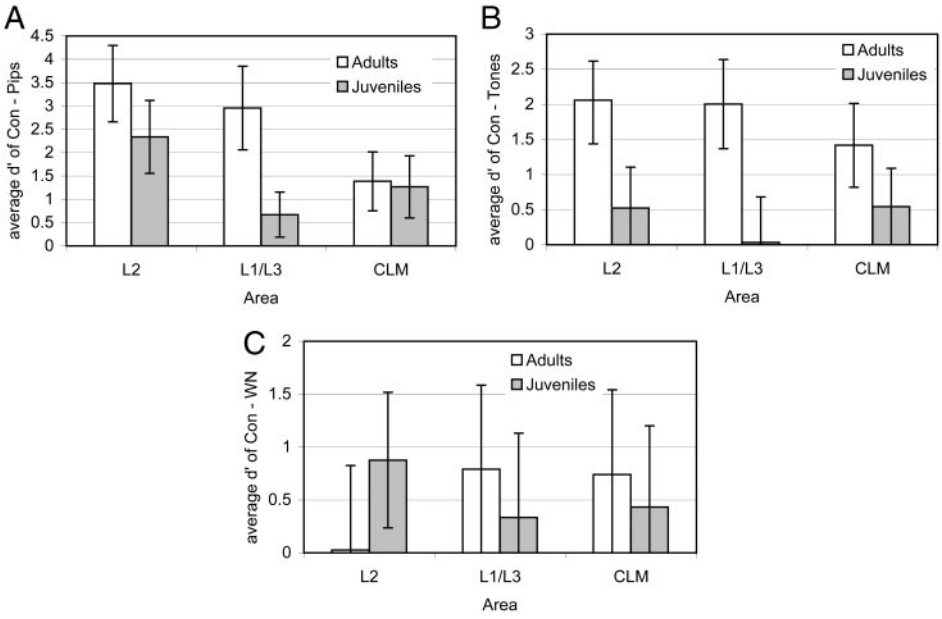



FIG. 6. Comparison of juvenile and adult mean d′ values across areas (L2, L1/L3, and CLM) for all the Con–Synthetic comparisons: *A*: Con–Pips. *B*: Con–Tones. *C*: Con–WN. L1/L3 subareas show the greatest developmental effect. Error bars represent 2 SEs.

2004). The information theoretic analysis requires a substantial amount of neural data and thus was not performed here. In summary, even though there is some selectivity for natural sounds or conspecific song in juveniles, this selectivity is not fully formed in field L and in particular in secondary subregions L1 and L3.

*Responses to the tutor song*

The tutor song did not seem to be represented preferentially in field L (L2 and L1/L3 subregions) of juveniles: Con and Tutor elicited similar response strengths, as shown by the similarity in their z-scores [$F(1,132) = 0.04$, $P = 0.83$] (see Fig. 7*A*). The average juvenile z-scores are once again about half (44%) that of the adults, suggesting that the juvenile auditory telencephalon does not respond as robustly to these behaviorally relevant natural stimuli as it does in adults. Indeed, a one-way ANOVA shows a difference in z-scores (pooled for the two stimuli) between adults and juveniles [$F(1,467) = 23.17$, $P < 0.0001$]. We further examined the possibility that juvenile neurons were selectively responsive to Tutor song in pairwise comparisons. Figure 7*B*, however, indicates that the mean of the cumulative probability distribution of d′ values for the juvenile Tutor–Con comparison is not different from zero, and that this distribution in juveniles overlaps a great deal with that of the adults. These findings are summarized in Fig. 7*C*, showing that on average there is no neural discriminability between the responses strengths for

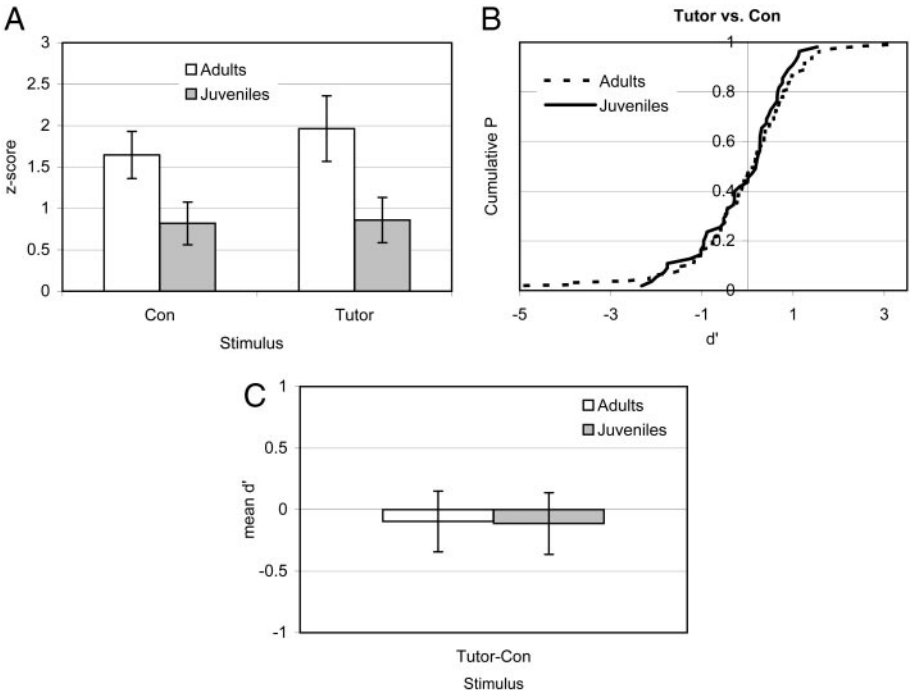



FIG. 7. Comparison of juvenile and adult responses to conspecific and bird's tutor's songs (Tutor) in field L (includes L2 and L1/L3 subregions). *A*: mean z-scores show no difference between Con and Tutor in the case of the juveniles. Average juvenile z-scores for Con and Tutor are significantly different from the corresponding adult average z-scores. *B*: cumulative distribution of the d′ values for the Tutor–Con comparison for both juveniles and adults. *C*: mean d′ measure for the Tutor–Con comparison shows no preference for Con or Tutor in the case of juveniles or adults. Error bars (in *A* and *C*) represent 2 SEs.

Tutor and Con (for either single units, multiunits, or combined-unit analysis) at this stage (combined units data reported here: mean d′ for Tutor–Con for juveniles = −0.11, $t = -0.91$, and $P = 0.37$), a finding that is similar to that of adults. Similar results are found in CLM: developing CLM neurons were also not able to discriminate Tutor from Con (combined units for CLM: mean d′ for Tutor–Con for juveniles = 0.21, $t = 1.20$, and $P = 0.26$).

### *Selectivity for the natural temporal structure of song*

We also tested the role of the temporal structure of syllables on the primary auditory telencephalon. Previous studies of the song system characterized selectivity for the bird's own song (BOS) in several ways, including *1*) comparing the response to BOS with the response elicited by the BOS played in reverse and *2*) comparing responses to BOS versus reverse-ordered BOS, which assesses selectivity for naturally occurring temporal combinations of syllables. Both temporal manipulations preserve the overall power in the song as well as the same overall modulation spectrum. We were interested in seeing what role such manipulations had on the selectivity of the young field L neurons (L2 and L1/L3 regions). As in adults, we found a small but significant selectivity in juveniles for the forward version of the song compared with its reverse, as opposed to the more "natural" reverse order [Con–CRev: mean d′ (for both single and multiunits combined) = 0.27, $t = 2.65$, $P = 0.01$; Con–CRO: mean d′ = 0.30, $t = 3.20$, $P = 0.002$; Tutor–TRev: mean d′ = 0.24, $t = 2.33$, $P = 0.02$; Tutor–TRO: mean d′ = 0.07, $t = 0.93$, $P = 0.35$]. This small preference for forward song is also seen in the adult data, as observed by the average d′ values for BOS over Rev (adult BOS–Rev: mean d′ = 0.25, $t = 43.39$, $P = 0.0008$; adult BOS–Revorder: mean d′ = 0.06, $t = 1.25$, $P = 0.21$). We interpret the selectivity for the forward sound not as a selectivity for a particular song such as the BOS (as it is used in the song system literature) but as a selectivity for natural sounds versus matched synthetic sounds, in this case the reverse songs.

### *Do responsivity and selectivity coemerge?*

We next asked whether the observed limited selectivity for Conspecific song versus the matched synthetic sounds in juveniles was correlated with the observed low responsivity. To show the relationship between responsivity and selectivity, we plotted juvenile and adult stimulus-excited sites' z-scores for Con and their corresponding d′ values for Con–Pips for all of our data (see Fig. 8*A*) and for subfields L1 and L3 (Fig. 8*B*), subregions where we found the largest difference between adults and juveniles. We found a positive correlation between the two metrics in both cases (all of the data: adult $R^2 = 0.64$, $P < 0.0001$; juvenile $R^2 = 0.81$, $P < 0.0001$; subfields L1/L3 only: adult $R^2 = 0.52$, $P < 0.0001$; juvenile $R^2 = 0.82$, $P < 0.0001$). Figure 8*B* reveals that juveniles' z-scores and d′ values from L1 and L3 cluster toward the smaller end of the scale, with a great deal of overlap with the z-scores and d′ values from the same subregions in adults, although the adult L1 and L3 neurons have a larger spread toward the positive end of the scale. Thus those cells that develop stronger responses also become more selective. To quantify this finding, we performed biased random sampling of the adult z-score distribution for L1 and L3 in response to Con to obtain subsamples with the same distribution as the juvenile Con z-score for the same subregions (see Fig. 8*C*). We then calculated the average selectivity for these subareas, using the d′ measure, for this subset of the adult data. Figure 8*D* shows that the new resampled adult d′ distribution is not statistically different from that of the juveniles' (mean d′ for resampled adult distribution = 1.43, SE = 0.44; mean d′ for juveniles = 1.19, SE = 0.30), thereby suggesting that responsivity and selectivity are properties that mature together in primary auditory forebrain development.

### *Coefficient of variation analysis*

Cross-trial CV values were computed for both adult and juvenile auditory forebrain neuronal responses (both field L and CLM combined) to Con (adults: average CV = 0.63, SE = 0.02; juveniles: average CV = 0.81, SE = 0.04). This result illustrates that the juvenile auditory telencephalon tended to be less reliable in its responses than its adult counterpart. We performed the CV analysis across subareas as well and found that both adult and juvenile CLM neurons are extremely variable and not statistically different from each other (CLM adults: average CV = 0.95, SE = 0.08; CLM juveniles: average CV = 0.91, SE = 0.09), which could partly explain the low d′ values for this nucleus. The variability of subregion L2 in juveniles, although statistically different from its adult counterpart, was more similar to the variability found in the adults (L2 adults: average CV = 0.55, SE = 0.03; L2 juveniles: average CV = 0.69, SE = 0.05) compared with the rest of the developing field L, which was extremely variable and also statistically different from the adult L1 and L3 subregions (L1/L3 adults: average CV = 0.56, SE = 0.04; L1/L3 juveniles: average CV = 0.91, SE = 0.07).

## DISCUSSION

We studied the response selectivity of auditory forebrain neurons of field L and CLM in 35-day-old urethane-anesthetized zebra finches and compared it to responses of adult neurons to the same stimuli (Grace et al. 2003). We found that auditory responses overall in field L in juveniles were not as robust as those found in the adults. Responses were smaller and showed greater variability. Selectivity for song over matched synthetic sounds was limited in 35-day-old juveniles; adultlike selectivity was observed in CLM but was reduced in field L and in particular in the secondary subregions L1 and L3. We found that selectivity for song and general auditory responsiveness coemerge during ontogeny in the secondary subregions of field L. Finally, no neural selectivity for the tutor song in either field L or CLM was observed. We subsequently discuss potential mechanisms underlying these physiological changes and the relevance of these results to song recognition and memory.

### *Immature responsivity and limited selectivity for song in field L*

The lower responsiveness (lower z-scores) and the higher variability in mean firing rates (higher CVs) imply that the overall development of the primary auditory telencephalon is still incomplete at 35 days of age. These results are consistent

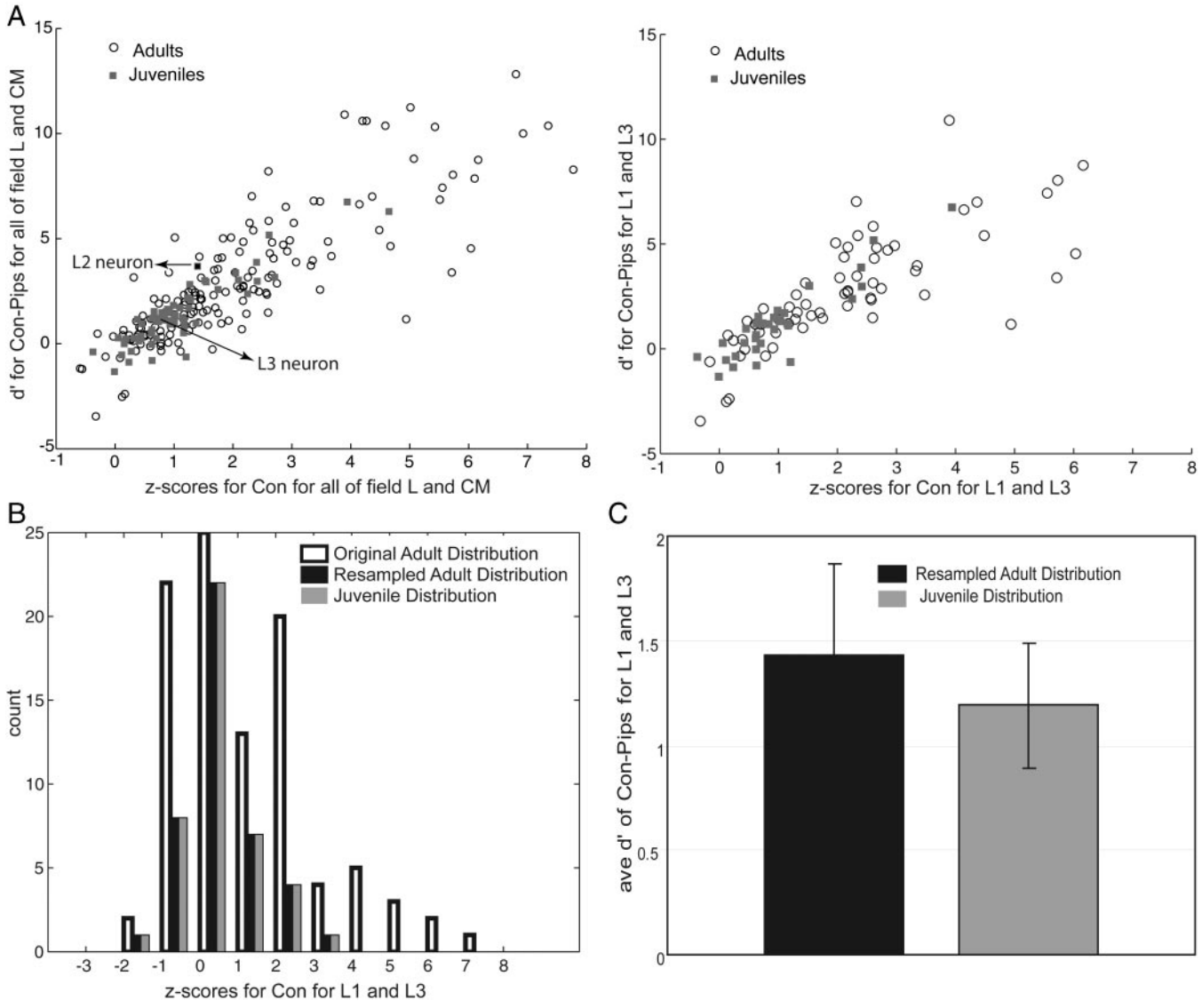



FIG. 8. Comparing the selectivity of juvenile and adult neurons with similar responsivity. *A*: correlation of z-scores of stimulus-excited sites for Con and their corresponding d′ values for the Con–Pips comparison for both adults and juveniles for field L and CLM. Two example cells whose responses were illustrated in Fig. 3 are indicated with arrows to show their position in this distribution. *B*: correlation of z-scores of stimulus-excited sites for Con and their corresponding d′ values for the Con–Pips comparison for both adults and juveniles restricted to subfields L1 and L3. *C*: count histograms of the z-scores in subfields L1 and L3 for conspecific song in adults (original adult distribution), in a biased sample of the adult data (resampled adult distribution) and in juveniles. Resampled adult z-score distribution was designed to match the juvenile z-score distribution (see METHODS). *D*: average d′ for Con vs. Pips of the adult resampled distribution (resampled to have similar z-scores to those of juveniles) is not significantly different from the d′ of the Con–Pips comparison in juveniles, suggesting that responsivity and selectivity coemerge during ontogeny in the secondary subregions of field L.

with other studies in both birds and mammals. In the auditory system of zebra finches, Gehr et al. (2000) showed that field L of 30-day-old finches has only three functional areas compared with six in adults. In juvenile rats, tuning curves in primary auditory cortex (A1) are broader in younger animals. In older rats, both tuning and tonotopy are refined (Zhang et al. 2001, 2002).

Selectivity for conspecific song relative to our matched synthetic sounds was also limited at this developmental phase in field L and in particular in subareas L1 and L3: in contrast to the adult, conspecific song elicited responses similar to those to matched complex tones and only partial selectivity was observed for song over simple tones. This limited selectivity in juveniles is reminiscent of the findings in the song system, where the tuning for the bird's own song emerges between 30 days and adulthood, during the vocal learning phase (Doupe 1997; Nick and Konishi 2005b; Solis and Doupe 1997). We also found that overall auditory responsiveness and neural selectivity for natural sounds versus synthetic sounds are correlated in adult birds: the adult neurons in these secondary subfields that are less responsive are also less selective and more like those found in the juvenile.

### *Maturation or experience-dependent plasticity?*

Although our experiments did not address developmental mechanisms, we believe both maturation and/or experience-based changes could support our observations. For example, the small degree of selectivity for song that is observed in the extrathalamo-recipient subregions of the primary auditory fore-

brain in juveniles could simply be reinforced, with synaptic mechanisms such as winner-take-all (Shun-Ichi and Arbib 1977), to yield the higher responsivity and selectivity observed in adults. It is also possible that experience-dependent plasticity plays a role in shaping the selectivity of auditory forebrain neurons for natural sounds, which are behaviorally reinforced. Developmental plasticity was previously demonstrated in the song system of zebra finches, where auditory experience is important in shaping selectivity for the bird's own song (Doupe 1997; Solis and Doupe 1997, 1999; Volman 1993). In addition, depriving juvenile male birds of normal auditory experience attenuated pruning of spines in the anterior forebrain pathway (Wallhauser-Franke et al. 1995) and delayed the emergence of topographic specificity within the lMAN core to RA circuit, a major anterior forebrain pathway implicated in song learning (Iyengar and Bottjer 2002).

Although such functional and structural experience-dependent changes were demonstrated in the song system (for review see Bottjer 2004), developmental plasticity studies of the auditory system have only begun. A recent neurophysiological study of the primary auditory forebrain in acoustically deprived starlings (Cousillas et al. 2004) is consistent with an important function for acoustic experience in shaping field L: a much larger number of adult field L neurons in these birds respond nonselectively to multiple stimuli than do control birds. Moreover, the deprived starlings did not have normal dorsoventral tonotopy in field L. In addition, social deprivation in starlings was recently shown to influence neural responsiveness by reducing the number of auditory responsive sites and selectivity in field L (Cousillas et al. 2006). Gentner and Margoliash (2003) found units in the CM of starlings to be selective for familiar songs learned during operant conditioning and thus this region appears to be influenced by the birds' perceptual history.

### *Selectivity differences between L2 and L1/L3*

The differences in selectivity and responsivity between juveniles and adults were smaller in the L2 region than in the L1/L3 regions. L2 is the major recipient of input from the thalamic nucleus ovoidalis (Vates et al. 1996) and the subregion L2a, in particular, can be thought to be analagous to layer 4 of the auditory cortex (Wild et al. 1993). Considering the thalamo-recipient region L2 to be low level relative to the other subregions (L1 and L3), our results are in accord with other neurodevelopmental processes where lower-level areas develop before high-level or more specialized areas (the contradictory adultlike selectivity found in the secondary CLM region is discussed in detail later). Similar findings were seen in the visual system, where the proportion of cells in the primary visual cortex that respond to visual stimuli increases with age, and more unresponsive cells are found in layers 2, 3, and 5 in young animals than in the layers that receive direct input from the visual thalamus (Albus and Wolf 1984). This hierarchical developmental processing in songbirds in particular is illustrated in Fig. 9.

The developmental differences between the thalamo-recipient subregion L2 and the secondary subregions of L1 and L3 raise the question of whether the connections between these subfields are fully developed at this age or whether greater plasticity in these intrafield circuits is mediated by strengthening of inhibitory synapses. To date, there are no studies of the auditory system in songbirds that have addressed whether these local circuits are developed at this stage. Immediate early gene (IEG) studies in adult zebra finches showed that the zenk gene is induced in response to hearing conspecific song in all of the auditory telencephalon (including L1, L3, CLM, and NCM) except in the subfield L2 (Mello and Clayton 1994). Zenk expression was proposed to be a marker of experience-dependent plasticity in songbirds and was linked to neuronal plasticity in mammalian studies (for review see Mello et al. 2004). Thus it is possible that L1 and L3 are either more plastic or plastic for a longer period of time than subfield L2. Similar findings regarding hierarchical development of plasticity across layers of primary sensory cortical areas were previously reported in mammals (Polley et al. 2004).

### *Development of perceptual behavior and the selectivity for natural sounds in extrathalamo-recipient field L*

Twenty-five-day-old male and female zebra finches do not engage at all in song-discrimination tests and, by 35 days of age, only about a third to a half of the experimental subjects participate in song-discrimination tests. However, by adulthood, all birds respond to the tests and are able to discriminate their father's song from that of another conspecific (Clayton 1988). Additionally, young birds raised by their father until day 25 and then isolated with siblings until sexual maturity do not show preference for father's song over another conspecific's song. By 35 days, however, zebra finches are beginning to show individual conspecific recognition and this recognition continues into adulthood (Clayton 1988; Miller 1979). Moreover, young male zebra finches can learn to produce a good copy of a tutor song to which they have been exposed until they are 35 days old (Böhner 1990). These behavioral observations taken together suggest that the auditory areas responsible for familiar conspecific song recognition should be beginning to mature by 35 days posthatch and that this auditory development should reach maturity by adulthood. The intermediate selectivity in subareas L1 and L3 at this developmental stage correlates with the start of individual conspecific song recognition or song memory in 35-day-old birds, which then continues to develop into adulthood. We believe that additional perceptual and vocal experience may be crucial for the emergence of natural sound selectivity in these secondary auditory areas. In parallel, the neural development of selective responses to natural sounds in these secondary subareas of field L could facilitate the improvement in behavioral recognition of individual conspecific songs as birds mature.

A somewhat contradictory result is the adultlike selectivity found in the higher-level auditory area of CLM. This result echoes the neurophysiological responses found in another secondary auditory area that was previously implicated in the representation of song, NCM (Stripling et al. 1997). In NCM, adultlike neurophysiological responses to conspecific song were reported in birds as young as 20 days (although the habituation rate to conspecific song is slower in juveniles than in adults), with adult IEG responses starting to be seen at 30 days and continuing into adulthood (Stripling et al. 2001). It is feasible that the adultlike neurophysiological response properties in these secondary areas are inherited from L2 because direct connections between L2 and both NCM and CLM were

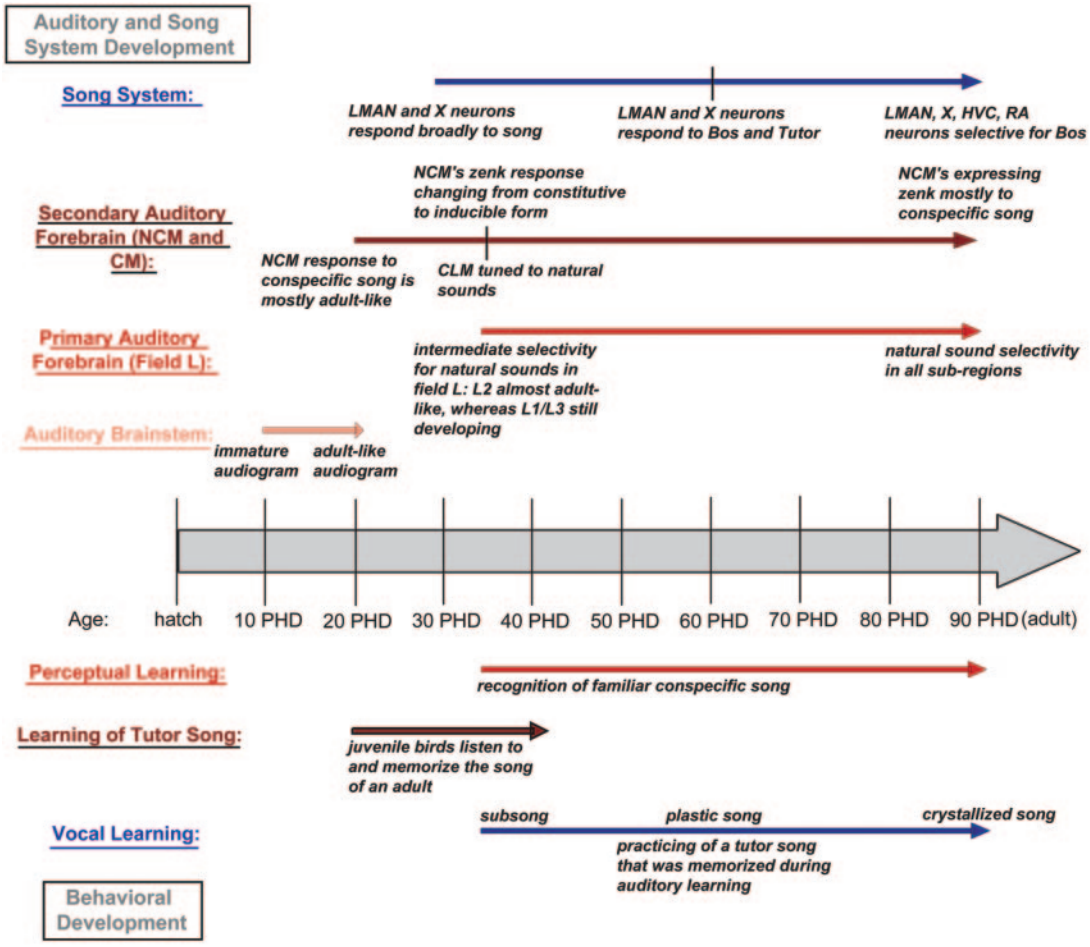



FIG. 9. Working model of the auditory system's (in red) developmental timeline in relation to the song system's (in blue) development.

reported. However, it is also possible—and we believe more probable (especially given the highly variable firing that we observed in CLM)—that the neurophysiological properties in CLM change during development but that these changes were not detected in our experiments: CLM receives strong projections from L3 and L1 (Vates et al. 1996), which as we have shown here change during development. Further neurophysiological recordings in awake and behaving young and adult birds will therefore be needed to fully understand the development of these areas.

*No selective tuning for the behaviorally relevant tutor song*

A particularly behaviorally relevant case of neural plasticity in birdsong is the storage of the tutor template, a special neural representation of the sound of the tutor song that is thought to guide vocal learning (Adret 2004; Theunissen et al. 2004b). Auditory forebrain neurons in adult field L and CLM show no neurophysiological evidence of selectivity for the bird's own song or tutor song (Amin et al. 2004). We tested whether tutor song selectivity transiently exists in field L and CLM of young birds, at a stage when they have memorized the tutor's song and require the memory for future template matching during singing (Eales 1985). Our results, consistent with those of Gehr et al. (2000), suggest that the tutor song is not neurally discriminated by differences in mean firing in these two auditory forebrain areas. A favored neural representation for tutor song could still be occurring in field L or CLM, although it would have to involve either a population of neurons that was not recorded in these experiments or a neural representation not based on mean firing rates. Also, it should be noted that we did not systematically use the tutor song as a search stimulus and thus it is possible that we missed recording from potential sites highly selective for the tutor's song. In addition, we did not record the tested birds' subsong vocalizations and so we cannot directly test other hypotheses about tutor tuning that relate to vocal behavior: for instance, perhaps only the best and most faithful vocal learners have tutor-selective cells in large numbers.

In contrast with the results that the sampled auditory forebrain neurons were not tutor selective, previous IEG expression studies suggested that secondary auditory regions of NCM and medial CM (CMM) in adults are modulated by early auditory experience of the tutor song. Zenk and c-fos responses in NCM to the tutor song were shown to be correlated with the strength of song learning in adult males (Bolhuis et al. 2000,2001; Terpstra et al. 2004). Similarly, increased CMM zenk responses to the tutor song were observed in adult females who also exhibited behavioral preferences for these songs (Terpstra et al. 2006). In line with these studies, a recent electrophysiological study in adult males also suggested a special representation for the tutor song in NCM (Phan et al. 2006): the slope of the neural adaptation to the tutor song was

shallower than that for novel songs and characteristic of familiar sounds even though the exposure to the tutor song had occurred relatively long ago. Moreover, the strength of this effect was correlated with how well the tutor song was copied.

We did not find tutor-selective responses in CLM but it should be noted that our neurophysiological studies were conducted on males (both young and adult) and in the lateral region of CM, whereas the IEG effect reported in Terpstra et al. (2006) study was in CMM and in adult females. Thus the mystery of the neural representation for the tutor template remains: this song memory could be in NCM, in a more lateral area of the auditory nidopallium, in CMM, in song nuclei (Nick and Konishi 2005a; Solis and Doupe 1999), or in more peripheral auditory areas. Additionally, the neural representation of the tutor song in young birds could reside in a sparse and distributed form that is not readily accessible by single-unit recordings.

In summary, we have shown that general responsivity and specific tuning for conspecific song emerges during late development in the primary auditory forebrain area field L and, in particular, in subregions L1 and L3. In addition, we did not find a marker of tutor exposure in either field L or CLM at this developmental stage. If the observed neural tuning for conspecific-like natural sounds is a correlate of the specific auditory computations that are required to recognize complex acoustical features in song, then the gradual development of this selective tuning could underlie the emergence of the individual song discrimination observed behaviorally. Additional acoustical experience, including that of a bird's own song and of that of others, could be critical for this development. Thus conducting behavioral and neurophysiological experiments in acoustically deprived birds will be crucial for understanding how experience and maturation interact to permit the emergence of the complex responses observed at these highest levels of auditory processing (Cousillas et al. 2004).

## ACKNOWLEDGMENTS

We thank Dr. Ed Rubel for demonstrating how to record auditory brain stem responses. We also thank S. M. Woolley for reading a previous version of this manuscript.

## GRANTS

This work was supported by National Institutes of Health Grants MH-059189 to F. E. Theunissen and NS-34835 and MH-55987 to A. J. Doupe.

## REFERENCES

**Adret P.** In search of the song template. *Ann NY Acad Sci* 1016: 303–324, 2004.

**Albus K, Wolf W.** Early postnatal development of neuronal function in the kitten's visual cortex: a laminar analysis. *J Physiol* 348: 153–185, 1984.

**Amin N, Grace JA, Theunissen FE.** Neural response to bird's own song and tutor song in the zebra finch field L and caudal mesopallium. *J Comp Physiol A Sens Neural Behav Physiol* 190: 469–489, 2004.

**Appeltants D, Gentner TQ, Hulse SH, Balthazart J, Ball GF.** The effect of auditory distractors on song discrimination in male canaries (*Serinus canaria*). *Behav Process* 69: 331–341, 2005.

**Aubin T, Jouventin P, Hildebrand C.** Penguins use the two-voice system to recognize each other. *Proc Biol Sci* 267: 1081–1087, 2000.

**Böhner J.** Early acquisition of song in the zebra finch, *Taneiopygia guttata. Anim Behav* 39: 369–374, 1990.

**Bolhuis JJ, Gahr M.** Neural mechanisms of birdsong memory. *Nat Rev Neurosci* 7: 347–357, 2006.

**Bolhuis JJ, Hetebrij E, Den Boer-Visser AM, De Groot JH, Zijlstra GG.** Localized immediate early gene expression related to the strength of song learning in socially reared zebra finches. *Eur J Neurosci* 13: 2165–2170, 2001.

**Bolhuis JJ, Zijlstra GG, den Boer-Visser AM, Van Der Zee EA.** Localized neuronal activation in the zebra finch brain is related to the strength of song learning. *Proc Natl Acad Sci USA* 97: 2282–2285, 2000.

**Bottjer SW.** Developmental regulation of basal ganglia circuitry during the sensitive period for vocal learning in songbirds. *Ann NY Acad Sci* 1016: 395–415, 2004.

**Brittan-Powell EF, Dooling RJ, Gleich O.** Auditory brainstem responses in adult budgerigars (*Melopsittacus undulatus*). *J Acoust Soc Am.* 112: 999–1008, 2002.

**Capsius B, Leppelsack HJ.** Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. *Hear Res* 96: 59–70, 1996.

**Cardin JA, Schmidt MF.** Auditory responses in multiple sensorimotor song system nuclei are co-modulated by behavioral state. *J Neurophysiol* 91: 2148–2163, 2004.

**Catchpole CK, Slater PBJ.** *Bird Song Biological Themes and Variations*. Cambridge, UK: Cambridge Univ. Press, 1995.

**Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS.** Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. *Proc Natl Acad Sci USA* 92: 3406–3410, 1995.

**Clayton NS.** Song discrimination learning in zebra finches. *Anim Behav* 36: 1016–1024, 1998.

**Cousillas H, George I, Mathelier M, Richard JP, Henry L, Hausberger M.** Social experience influences the development of a central auditory area. *Naturwissenschaften* 93: 588–596, 2006.

**Cousillas H, Richard JP, Mathelier M, Henry L, George I, Hausberger M.** Experience-dependent neuronal specialization and functional organization in the central auditory area of a songbird. *Eur J Neurosci* 19: 3343–3352, 2004.

**Dooling RJ, Brown SD, Klump GM, Okanoya K.** Auditory perception of conspecific and heterospecific vocalizations in birds: evidence for special processes. *J Comp Psychol* 106: 20–28, 1992.

**Doupe AJ.** Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J Neurosci* 17: 1147–1167, 1997.

**Eales LA.** Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Anim Behav* 33: 1293–1300, 1985.

**Fortune ES, Margoliash D.** Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (*Taenopygia guttata*). *J Comp Neurol* 325: 388–404, 1992.

**Gehr DD, Hofer SB, Marquardt D, Leppelsack H.** Functional changes in field L complex during song development of juvenile male zebra finches. *Brain Res Dev Brain Res* 125: 153–165, 2000.

**Gentner TQ, Margoliash D.** Neuronal populations and single cells representing learned auditory objects. *Nature* 424: 669–674, 2003.

**George I, Vernier B, Richard JP, Hausberger M, Cousillas H.** Hemispheric specialization in the primary auditory area of awake and anesthetized starlings (*Sturnus vulgaris*). *Behav Neurosci* 118: 597–610, 2004.

**Grace JA, Amin N, Singh NC, Theunissen FE.** Selectivity for conspecific song in the zebra finch auditory forebrain. *J Neurophysiol* 89: 472–487, 2003.

**Heinrich JE, Singh TD, Sohrabji F, Nordeen KW, Nordeen EJ.** Developmental and hormonal regulation of NR2A mRNA in forebrain regions controlling avian vocal learning. *J Neurobiol* 51: 149–159, 2002.

**Hile AG, Plummer TK, Striedter GF.** Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Anim Behav* 59: 1209–1218, 2000.

**Hsu A, Woolley SM, Fremouw TE, Theunissen FE.** Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *J Neurosci* 24: 9201–9211, 2004.

**Iyengar S, Bottjer SW.** The role of auditory experience in the formation of neural circuits underlying vocal learning in zebra finches. *J Neurosci* 22: 946–958, 2002.

**Iyengar S, Wiswanathan SS, Bottjer SW.** Development of topography within song control circuitry of zebra finches during the sensitive period for song learning. *J Neurosci* 19: 6037–6057, 1999.

**Janata P, Margoliash D.** Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *J Neurosci* 19: 5108–5118, 1999.

**Jarvis ED, Nottebohm F.** Motor driven gene expression. *Proc Natl Acad Sci USA* 94: 4097–4102, 1997.

**Leonard ML, Horn AG.** Ambient noise and the design of begging signals. *Proc Biol Sci* 272: 651–656, 2005.

**Leppelsack HJ, Vogt M.** Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. *J Comp Neurol* 107: 263–274, 1976.

**Lewicki MS, Arthur BJ.** Hierarchical organization of auditory temporal context sensitivity. *J Neurosci* 16: 6987–6998, 1996.

**MacDougall-Shackleton SA, Hulse SH, Ball GF.** Neural bases of song preferences in female zebra finches (*Taeniopygia guttata*). *Neuroreport* 9: 3047–3052, 1988.

**Margoliash D.** Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *J Neurosci* 6: 1643–1661, 1986.

**Marler P.** Bird calls: their potential for behavioral neurobiology. *Ann NY Acad Sci* 1016: 31–44, 2004.

**Marler P, Peters S.** Long-term storage of learned birdsong prior to production. *Anim Behav* 30: 479–482, 1982.

**Marshall-Bal L, Slater PJ.** Duet singing and repertoire use in threat signalling of individuals and pairs. *Proc Biol Sci* 271: S440–S443, 2004.

**Mello CV, Clayton DF.** Song-induced ZENK gene expression in auditory pathways of songbird brain and its relation to the song control system. *J Neurosci* 14: 6652–6666, 1994.

**Mello CV, Ribeiro S.** ZENK protein regulation by song in the brain of songbirds. *J Comp Neurol* 393: 426–438, 1988.

**Mello CV, Velho TA, Pinaud R.** Song-induced gene expression: a window on song auditory processing and perception. *Ann NY Acad Sci* 1016: 263–281, 2004.

**Mello CV, Vicario DS, Clayton DF.** Song presentation induces gene expression in the songbird forebrain. *Proc Natl Acad Sci USA* 89: 6818–6822, 1992.

**Miller DB.** Long-term recognition of father's song by female zebra finches. *Nature* 280: 389–391, 1979.

**Muller CM, Leppelsack HJ.** Feature extraction and tonotopic organization in the avian auditory forebrain. *Exp Brain Res* 59: 587–599, 1985.

**Nick TA, Konishi M.** Neural song preference during vocal learning in the zebra finch depends on age and state. *J Neurobiol* 62: 231–242, 2005a.

**Nick TA, Konishi M.** Neural auditory selectivity develops in parallel with song. *J Neurobiol* 62: 469–481, 2005b.

**Nordeen KW, Nordeen EJ.** Synaptic and molecular mechanisms regulating plasticity during early learning. *Ann NY Acad Sci* 1016: 416–437, 2004.

**Nottebohm F, Stokes TM, Leonard CM.** Central control of song in the canary, *Serinus canarius. J Comp Neurol* 165: 457–486, 1976.

**Phan ML, Pytte CL, Vicario DS.** Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proc Natl Acad Sci USA* 103: 1088–1093, 2006.

**Polley DB, Kvasnak E, Frostig RD.** Naturalistic experience transforms sensory maps in the adult cortex of caged animals. *Nature* 429: 67–71, 2004.

**Ribeiro S, Cecchi GA, Magnasco MO, Mello CV.** Toward a song code evidence for a syllabic representation in the canary brain. *Neuron* 21: 359–371, 1988.

**Riebel K.** Early exposure leads to repeatable preferences for male song in female zebra finches. *Proc R Soc Lond B Biol Sci* 267: 2553–2558, 2000.

**Riebel K, Smallegange IM, Terpstra NJ, Bolhuis JJ.** Sexual equality in zebra finch song preference evidence for a dissociation between song recognition and production learning. *Proc R Soc Lond B Biol Sci* 269: 729–733, 2002.

**Sherman P, Reeve H, Pfennig D.** Recognition systems. In: *Behavioral Ecology*, edited by Krebs J, Davies NB. Malden, MA: Blackwell Science, 1997.

**Shun-Ichi A, Arbib MA.** Competition and cooperation in neural nets. In: *Systems Neuroscience*, edited by Metzler J. New York: Academic Press, 1977, p. 119–165.

**Singh NC, Theunissen FE.** Modulation spectra of natural sounds and ethological theories of auditory processing. *J Am Stat Assoc* 114: 3394–3411, 2003.

**Slater PJB, Eales LA, Clayton NS.** Song learning in zebra finches (*Taeniopygia guttata*): progress and prospects. *Adv Study Behav* 18: 1–33, 1988.

**Solis MM, Doupe AJ.** Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. *J Neurosci* 17: 6447–6462, 1997.

**Solis MM, Doupe AJ.** Contributions of tutor and bird's own song experience to neural selectivity in the songbird anterior forebrain. *J Neurosci* 19: 4559–4584, 1999.

**Stripling R, Kruse AA, Clayton DF.** Development of song responses in the zebra finch caudomedial neostriatum: role of genomic and electrophysiological activities. *J Neurobiol* 48: 163–180, 2001.

**Stripling R, Volman SF, Clayton DF.** Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. *J Neurosci* 17: 3883–3893, 1997.

**Terpstra NJ, Bolhuis JJ, den Boer-Visser AM.** An analysis of the neural representation of birdsong memory. *J Neurosci* 24: 4971–4977, 2004.

**Terpstra NJ, Bolhuis JJ, Riebel K, van der Burg JM, den Boer-Visser AM.** Localized brain activation specific to auditory memory in a female songbird. *J Comp Neurol* 494: 784–791, 2006.

**Theunissen FE, Amin N, Shaevitz SS, Woolley SM, Fremouw T, Hauber ME.** Song selectivity in the song system and in the auditory forebrain. *Ann NY Acad Sci* 1016: 222–245, 2004b.

**Theunissen FE, Shaevitz SS.** Auditory processing of vocal sounds in birds. *Curr Opin Neurobiol* 16: 400–407, 2006.

**Theunissen FE, Woolley SM, Hsu A, Fremouw T.** Methods for the analysis of auditory processing in the brain. *Ann NY Acad Sci* 1016: 187–207, 2004a.

**Vates GE, Broome BM, Mello CV, Nottebohm F.** Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*Taenopygia guttata*). *J Comp Neurol* 366: 613–642, 1996.

**Volman SF.** Development of neural selectivity for birdsong during vocal learning. *J Neurosci* 13: 4737–4747, 1993.

**Wallhausser-Franke E, Nixdorf-Bergweiler BE, DeVoogd TJ.** Song isolation is associated with maintaining high spine frequencies on zebra finch lMAN neurons. *Neurobiol Learn Mem* 64: 25–35, 1995.

**Wild JM, Karten HJ, Frost BJ.** Connections of the auditory forebrain in the pigeon (*Columba livia*). *J Comp Neurol* 337: 32–62, 1993.

**Woolley SM, Fremouw TE, Hsu A, Theunissen FE.** Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat Neurosci* 8: 1371–1379, 2005.

**Woolley SM, Rubel EW.** Bengalese finches Lonchura striata domestica depend upon auditory feedback for the maintenance of adult song. *J Neurosci* 17: 6380–6390, 1997.

**Zevin JD, Seidenberg MS, Bottjer SW.** Limits on reacquisition of song in adult zebra finches exposed to white noise. *J Neurosci* 24: 5849–5862, 2004.

**Zhang LI, Bao S, Merzenich MM.** Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat Neurosci* 4: 1123–1130, 2001.

**Zhang LI, Bao S, Merzenich MM.** Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. *Proc Natl Acad Sci USA* 99: 2309–2314, 2002.